1 Trochę przypomnień i motywacji

2 Definicje i wyjście trochę poza nie

2.1 Przestrzeń wektorowa, liniowa niezależność, baza

2.1.1 Przestrzeń wektorowa

Def. Przestrzeń wektorowa

Przykłady

1. Przykład kanoniczny: $\mathbb{K}^n = \mathbb{K} \times \mathbb{K} \times \cdots \times \mathbb{K}$ (n razy) – iloczyn kartezjański n egzemplarzy ciała $\mathbb{K}$. (W dalszym ciągu będziemy używać jedynie ciał $\mathbb{R}$ i $\mathbb{C}$.) Dodawanie wektorów i mnożenie wektora przez liczbę definiujemy analogicznie jak w przypadku trójwymiarowym:

$$\text{Dla } v = \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix}, \quad w = \begin{bmatrix} w^1 \\ w^2 \\ \vdots \\ w^n \end{bmatrix}$$

$$\text{mamy}: \quad v + w = \begin{bmatrix} v^1 + w^1 \\ v^2 + w^2 \\ \vdots \\ v^n + w^n \end{bmatrix} \text{oraz} \quad \alpha v = \begin{bmatrix} \alpha v^1 \\ \alpha v^2 \\ \vdots \\ \alpha v^n \end{bmatrix}$$

   Uwaga: v^k oznacza tu $k-$tą składową wektora v, a nie $k-$tą potęgę v ! Okazuje się, że numerowanie warto prowadzić na dwa sposoby: Dla jednego rodzaju obiektów indeksy (numery) piszemy *z prawej strony na górze* (tu robimy to dla składowych wektorów), a dla drugiego (numery elementów bazy) piszemy *na dole*. Niedługo się okaże, dlaczego warto robić takie rozróżnienie.

2. $C([a, b])$ – przestrzeń funkcji ciągłych na odcinku $[a, b]$. Mając dane dwie funkcje $f, g \in C([a, b])$ (tzn. patrzymy na funkcje f, g jako na wektory), definiujemy ich sumę jako funkcję $f + g$, której wartość w punkcie x jest: $(f + g)(x) = f(x) + g(x)$; oraz mnożenie wektora f przez liczbę $\lambda \in \mathbb{R}$ jest funkcją λf, której wartość w punkcie x określamy jako: $(\lambda f)(x) = \lambda f(x)$.

3. $\mathbb{K}_n[\cdot]$ – przestrzeń wielomianów stopnia n o współczynnikach z ciała $\mathbb{K}$. Dodawanie wielomianów oraz mnożenie ich przez liczbę definiujemy analogicznie jak dla funkcji.

2.1.2 Kombinacja liniowa, liniowa niezależność

Def. *Kombinacją liniową* wektorów $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ z współczynnikami $\lambda_1, \lambda_2, \dots, \lambda_n$ nazywamy wektor

$$\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \cdots + \lambda_n \mathbf{x}_n$$

Def. Układ wektorów $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ nazywamy *liniowo niezależnym*, gdy równość

$$\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \cdots + \lambda_n \mathbf{x}_n = \mathbf{0}$$

zachodzi tylko dla współczynników *zerowych*: $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

Def. Mówimy, że przestrzeń wektorowa V ma *wymiar* n, gdy każdy układ $n+1$ wektorów z V jest liniowo zależny, zaś istnieją w V układy n wektorów liniowo niezależnych.

Wymiar przestrzeni V oznaczamy $\dim V$ (od ang. *dimension*).

2.1.3 Baza

Def. *Bazą* w przestrzeni wektorowej wymiaru n nazywamy układ n wektorów liniowo niezależnych.

Przykł. (kontynuacja powyższych)

1. Przestrzeń $\mathbb{K}^n$ ma wymiar n. Jako bazę można w niej wybrać $e_1, e_2, \ldots, e_n$:
$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$
(tzn. w wektorze e_k na $k-$tym miejscu mamy 1, pozostałe składowe są zerowe). Wymiar przestrzeni $\mathbb{K}^n$ wynosi n.

2. $\dim C([0,1]) = \infty$. Na tej konstatacji poprzestaniemy. Bazę można wybrać, ale wymaga to dużego wkładu z analizy, wychodzącego poza ramy tego wykładu.

3. $\dim \mathbb{K}_n[\cdot] = n+1$. Bo: Z tego co ongiś mówiliśmy o wielomianach, pamiętamy, że wielomian jest określony jednoznacznie przez podanie swoich współczynników (przy potęgach t od 0 do n – zakładamy, że mamy wielomiany od zmiennej t). Jako bazę można wybrać: $e_0 = 1, e_1 = t, e_2 = t^2, \ldots, e_n = t^n$.

Ale! W przykładach 1. i 3. bazy można powybierać inaczej. Weźmy np. $V = \mathbb{R}^3$. Mamy tu bazę standardową:
$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix};$$
ale w niczym nie gorszy jest wybór:
$$E_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$
Sprawdźmy, że jest to baza: Równość
$$\lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3 = \mathbf{0}$$
jest równoważna układowi równań:
$$\begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = 0 \\ \qquad\ \lambda_2 + \lambda_3 = 0 \\ \qquad\qquad\ \lambda_3 = 0 \end{cases}$$
który ma jedyne rozwiązanie
$$\lambda_1 = \lambda_2 = \lambda_3 = 0.$$

2.1.4 Rozkład wektora w bazie

Mamy proste ale ważne

Stw. Niech w przestrzeni wektorowej V będzie zadana baza $e = (e_1, e_2, \dots, e_n)$. Wtedy każdy wektor $v \in V$ ma jednoznaczny rozkład w bazie e:

$$v = v^1 e_1 + v^2 e_2 + \cdots + v^n e_n, \tag{1}$$

tzn. współczynniki $v^1, v^2, \dots, v^n$ są jednoznacznie wyznaczone.

Dow. Załózmy, że rozkład v nie jest jednoznaczny, tzn. że istnieje rozkład identyczny jak (1), tyle że z innymi współczynnikami:

$$v = s^1 e_1 + s^2 e_2 + \cdots + s^n e_n; \tag{2}$$

odejmijmy stronami (1) i (2). Mamy:

$$(s^1 - v^1)e_1 + (s^2 - v^2)e_2 + \cdots + (s^n - v^n)e_n = \mathbf{0}. \tag{3}$$

Ale jeżeli któryś ze współczynników s^k ma być różny od któregoś ze współczynników v^k, to $s^k - v^k \neq 0$, czyli jeśli któryś ze współczynników równości (3) jest niezerowy. A to oznacza sprzeczność z założeniem o liniowej niezależności e.

CBDO

Ozn. Często przydatne jest zaznaczanie, w jakiej bazie rozkładamy wektor. Jeśli rozkładamy wektor v w bazie e i otrzymujemy przy tym rozkład (1), to ten fakt oznaczamy jako:

$$[v]^e = \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix}; \tag{4}$$

jest to po prostu inny zapis postaci rozkładu (1), z jawnym zaznaczeniem, że składowe wektora v to współczynniki rozkładu w bazie e.

Wspominaliśmy, że w danej przestrzeni może być wiele różnych baz. Załóżmy, że mamy w danej przestrzeni V dwie bazy. Jak są powiązane współczynniki rozkładu jakiegoś wektora v w obu tych bazach?

W pełnej ogólności na to pytanie odpowiemy nieco później, po podaniu kilku faktów nt. odwzorowań liniowych. Na razie zauważmy, że mając współczynniki rozkładu w jednej bazie e, możemy uzyskać współczynniki rozkładu w innej bazie E przez *rozwiązanie układu równań.*

Przykł. Rozkład tego samego wektora w różnych bazach.

Niech $V = \mathbb{R}_2[\cdot]$. Jako pierwszą bazę weźmy bazę standardową e:

$$e_0 = 1, \quad e_1 = t, \quad e_2 = t^2 \tag{5}$$

Weźmy też drugą bazę E, zdefiniowaną jako:

$$E_0 = 1, \quad E_1 = 1 + t, \quad E_2 = 1 + t + t^2 \tag{6}$$

. Analogicznie jak w przykładzie sprzed dwóch stron, przekonujemy się, że E jest bazą.

Miejmy teraz zadany jakiś wektor $v \in V$; weźmy: $v \equiv v(t) = 3t^2 + 5t - 2$. Rozkład wektora v w bazie e jest natychmiastowy:

$$v = -2e_0 + 5e_1 + 3e_2. \tag{7}$$

Odwołując się do zapisu (1), mamy: $v = v^0 e_0 + v^1 e_1 + v^2 e_2$, czyli: $v^0 = -2$, $v^1 = 5$, $v^2 = 3$. Stosując sposób oznaczania (4), mamy:

$$[v]^e = \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix} \tag{8}$$

Rozłóżmy teraz wektor v w bazie E. Uczyńmy to na dwa sposoby.

• Pierwszy sposób. Oznaczmy współczynniki rozkładu w bazie E przez $\alpha^0, \alpha^1, \alpha^2$. Mamy więc: $v = \alpha^0 E_0 + \alpha^1 E_1 + \alpha^2 E_2$. Korzystając z definicji (6), mamy:

$$v = \alpha^0 \cdot 1 + \alpha^1(1+t) + \alpha^2(1+t+t^2) = (\alpha^0 + \alpha^1 + \alpha^2) \cdot 1 + (\alpha^1 + \alpha^2) \cdot t + \alpha^2 t^2,$$

co daje układ równań na współczynniki:

$$\begin{cases} \alpha^0 + \alpha^1 + \alpha^2 &= -2 \\ \alpha^1 + \alpha^2 &= 5 \\ \alpha^2 &= 3 \end{cases}$$

który łatwo rozwiązać, i otrzymujemy

$$\alpha^0 = -7, \quad \alpha^1 = 2, \quad \alpha^2 = 3.$$

Możemy zapisać otrzymany wynik jako

$$[v]^E = \begin{bmatrix} -7 \\ 2 \\ 3 \end{bmatrix} \tag{9}$$

• • Drugi sposób. Spójrzmy na definicję (6) bazy E jako na wyrażenie wektorów bazy E przez wektory bazy e:

$$E_0 = 1 = e_0; \quad E_1 = 1 + t = e_0 + e_1; \quad E_2 = 1 + t + t^2 = e_0 + e_1 + e_2.$$

Odwróćmy teraz tę zamianę zmiennych, tzn. wyraźmy wektory bazy e przez wektory bazy E. W powyższym przypadku łatwo ten układ rozwiązać. Mamy:

$$e_0 = E_0;$$

$$e_1 = E_1 - e_0 = E_1 - E_0;$$

$$e_2 = E_2 - e_0 - e_1 = E_2 - E_0 - (E_1 - E_0) = E_2 - E_1.$$

Skorzystajmy teraz z rozkładu (7) wektora v w bazie e; mamy:

$$\begin{aligned} v = -2e_0 + 5e_1 + 3e_2 &= -2 \cdot E_0 + 5(E_1 - E_0) + 3(E_2 - E_1) \\ &= 3E_2 + (5-3)E_1 + (-2-5)E_0 = 3E_2 + 2E_1 - 7E_0 \\ &= \alpha^0 E_0 + \alpha^1 E_1 + \alpha^2 E_2; \end{aligned}$$

czyli otrzymaliśmy $[v]^E$ – rozkład (9) wektora v na składowe w bazie E – to samo co w pierwszym sposobie.

2.2 Odwzorowanie liniowe

Def. Niech V, W – dwie przestrzenie wektorowe. Mówimy, że odwzorowanie $F : V \to W$ jest *liniowe*, gdy dla dowolnych wektorów $x, y \in V$ i dla dowolnego elementu ciała $\alpha \in \mathbb{K}$ spełnione są warunki:

$$F(x + y) = F(x) + F(y), \text{ (addytywność) } \quad F(\alpha v) = \alpha F(v) \text{ (jednorodność|)}. \tag{10}$$

Czasem te warunki zapisuje się w postaci jednego ('2 w 1')

$$\underset{x,y\in V}{\forall} \underset{\alpha,\beta\in\mathbb{K}}{\forall} : F(\alpha x + \beta y) = \alpha F(x) + \beta F(y). \tag{11}$$

Przykł.

1. $\mathbb{R}_n[\cdot] \to \mathbb{R}$: Branie wartości wielomianu w (zadanym) punkcie x.
2. $C([a, b]) \to \mathbb{R}$: $C([a, b]) \ni f \to \int_a^b f(x)\mathsf{d}x$
3. Kot Arnolda: $\mathbb{R}^2 \to \mathbb{R}^2$: $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} 2x_1 + x_2 \\ x_1 + x_2 \end{bmatrix}$

   • Zanim zaczniemy pokazywać liniowość, wprowadźmy pojęcie *macierzy* oraz *mnożenia macierzy przez wektor.*
   • Macierz ($m \times n$ – o m wierszach i n kolumnach): Tablica liczb: Dokładniej:

   Def. *Macierz* A rozmiaru $m \times n$ (o m wierszach i n kolumnach) i elementach z ciała $\mathbb{K}$ to tablica liczb postaci

   $$A = \begin{bmatrix} {a^1}_1 & \dots & {a^1}_n \\ \vdots & & \vdots \\ {a^m}_1 & \dots & {a^m}_n \end{bmatrix}, \tag{12}$$

   gdzie ${a^i}_j$ (*elementy macierzy*) są liczbami z ciała $\mathbb{K}$.

   • Indeksowanie macierzy.
   • Na wektor też można patrzeć jako na macierz.
   • Działanie macierzy na wektor (tzn. mnożenie macierzy przez wektor).
   • *Stw.* (na razie bardziej agitacyjne) Najogólniejsze odwzorowanie liniowe $F : V \to W$ to *mnożenie pewnej macierzy (macierz operatora F w bazach w V i W) przez wektor.*
   • Sprawdzenie że to jest liniowe

4. Różniczkowanie wielomianów
5. Obrót w $\mathbb{R}^3$ (wokół osi z)
6. Rzut

Def. Ker, Im

Przykł. Ker, Im przy różniczkowaniu

3 Macierze odwzorowań liniowych

Rozpatrzmy teraz zbiór *wszystkich możliwych* odwzorowań liniowych z V do W. Nazwiemy go $L(V,W)$.

Zbiorowi temu łatwo nadamy strukturę przestrzeni wektorowej – podobnie jak to było w przypadku zbioru funkcji ciągłych. Tu definicja będzie wyglądała tak:

$$(F_1+F_2)(x)=F_1(x)+F_2(x), \qquad (\alpha F)(x)=\alpha F(x)$$

Przykł. $L(R^1,R^1)$: $y=ax$ dla $x\in V=\mathbb{R}^1$, $y\in W=\mathbb{R}^1$. Odwzorowanie takie jest jednoznacznie wyznaczone przez jedną liczbę a. Suma dwóch odwzorowań: $M_a: y=ax$ oraz $M_b: y=bx$ to odwzorowanie $y=(a+b)x$.

Jak będzie w ogólnym przypadku, tzn. ile jest odwzorowań liniowych z V do W? Jak opisać (czy też: parametryzować) ich zbiór (będący – jak widzieliśmy – przestrzenią liniową)?

Przykład. Jeśli $V=W=\mathbb{R}$, to odwzorowanie liniowe $T: V\to W$ ma postać: $V\ni v\to w=\alpha v\in W$ ($\alpha\in\mathbb{R}$), tzn. odwzorowanie to jest po prostu mnożeniem przez liczbę. Odwzorowań liniowych z $\mathbb{R}^1$ do $\mathbb{R}^1$ jest więc tyle ile liczb α, a więc zbiór tych odwzorowań to $\mathbb{R}^1$: $L(V,W)\sim\mathbb{R}$.

W ogólnym przypadku mamy

Stwierdzenie. Jeśli $e=(e_1,\dots,e_n)$ – baza w V oraz $w_1,\dots,w_n$ – jakiekolwiek wektory z W, to istnieje dokładnie jedno $T\in L(V,W)$ takie, że $Te_1=w_1,\dots,Te_n=w_n$. Innymi słowy, odwzorowanie liniowe jest jednoznacznie wyznaczone przez swoje wartości na wektorach bazy.

Dowód: *Zdefiniujmy* działanie $T\in L(V,W)$ na wektorach bazy tak jak powyżej. Wtedy dla dowolnego wektora $v=\sum_{i=1}^n v^ie_i$ mamy

$$T(v)=T(\sum_{i=1}^n v^ie_i)=\sum_{i=1}^n v^iT(e_i)=\sum_{i=1}^n v^iw_i,$$

do daje szukany wzór na T (dokładniej, na $T(v)$). Pozostaje upewnić się, czy ten wzór określa odwzorowanie liniowe. Biorąc $v=\sum_{i=1}^n v^ie_i$ i $\tilde v=\sum_{i=1}^n \tilde v^ie_i$, mamy: $v+\tilde v=\sum_{i=1}^n(v^i+\tilde v^i)e_i$ oraz

$$T(v+\tilde v)=\sum_{i=1}^n(v^i+\tilde v^i)w_i=\sum_{i=1}^n v^iw_i+\sum_{i=1}^n\tilde v^iw_i=T(v)+T(\tilde v).$$

Podobnie sprawdza się jednorodność.

CBDO

Wniosek: Baza e w V ustala wzajemnie jednoznaczną odpowiedniość pomiędzy elementami $L(V,W)$ a ciągami $(w_1,\dots,w_n)$ wektorów z W. Innymi słowy, cała informacja o T jest zawarta w ciągu $[T]_e$:

$$[T]_e=(Te_1,\dots,Te_n).$$

$[T]_e$ jest ciągiem wektorów, a każdy z nich jest kolumienką liczb. Wyraźmy więc $[T]_e$ w bardziej konkretny sposób (jako zbiór liczb). Najsampierw, ustalmy bazę $f=f_1,\dots,f_m$ w przestrzeni W. Każdy z wektorów Te_j można zapisać $[Te_j]^f\in\mathbb{K}^m$. Zatem cała informacja o T jest zawarta w ciągu (dwuwskaźnikowym) liczb $[T]^f{}_e$:

$$[T]^f{}_e=([Te_1]^f,\dots,[Te_n]^f).$$

Oznaczając współrzędne wektora $[Te_j]^f$ przez $T^i{}_j$ (tu $i = 1, \dots, m$) możemy zapisać:

$$[Te_j]^f = \begin{bmatrix} T^1{}_j \\ \vdots \\ T^m{}_j \end{bmatrix} \tag{13}$$

a dla ciągu wektorów $[T]^f{}_e$

$$[T]^f{}_e = ([Te_1]^f, \dots, [Te_n]^f) = \left(\begin{bmatrix} T^1{}_1 \\ \vdots \\ T^m{}_1 \end{bmatrix}, \dots, \begin{bmatrix} T^1{}_n \\ \vdots \\ T^m{}_n \end{bmatrix} \right).$$

Ilość nawiasów w tym ostatnim wyrażeniu jest zdecydowanie za duża; zapiszmy to więc inaczej:

$$[T]^f{}_e = ([Te_1]^f, \dots, [Te_n]^f) = \begin{bmatrix} T^1{}_1 & \dots & T^1{}_n \\ \vdots & & \vdots \\ T^m{}_1 & \dots & T^m{}_n \end{bmatrix}. \tag{14}$$

To – po przypomnieniu sobie, co to jest macierz – w naturalny sposób prowadzi nas do definicji:

Def. *Macierz odwzorowania liniowego* $T \in L(V, W)$ w zadanych bazach: e w V oraz f w W, to tablica liczb (macierz) $[T]^f{}_e$ opisana wyżej.

Ustalenie baz w V oraz W zadaje wzajemnie jednoznaczną odpowiedniość pomiędzy elementami $L(V, W)$ oraz macierzami $m \times n$.

Podsumujmy związki pomiędzy wprowadzonymi wyżej trzema pojęciami: (T – odwzorowanie liniowe; $[Te_j]^f$ – macierz tego odwzorowania; $T^i{}_j$ – elementy tej macierzy.

Mając bazy e w V oraz f w W, elementy $T^i{}_j$ możemy liczyć na dwa sposoby:

1. **Sposób I** (transformacja wektorów bazy):

   $$Te_i = T^1{}_i f_1 + \dots + T^m{}_i f_m = \sum_{j=1}^{m} T^j{}_i f_j$$

   (rozwijamy w bazie f działania operatora T na poszczególne wektory bazy e).

   Jest to po prostu inna postać wzoru (13).

2. **Sposób II** (transformacja składowych): Jeśli $Tv = w$, gdzie $V \ni v = \sum_{i=1}^{n} v^i e_i$, $W \ni w = \sum_{j=1}^{m} w^j f_j$, to

   $$w^j = \sum_{i=1}^{n} T^j{}_i v^i \tag{15}$$

   Wynika to z poprzedniego sposobu przez proste przeliczenie:

   $$Tv = T(\sum_{i=1}^{n} v^i e_i) = \sum_{i=1}^{n} v^i T(e_i) = \sum_{i=1}^{n} v^i (\sum_{j=1}^{m} T^j{}_i f_j) = \sum_{j=1}^{m} (\sum_{i=1}^{n} T^j{}_i v^i) f_j$$

   $$= w = \sum_{j=1}^{m} w^j f_j$$

   i skoro tak, to mamy na składową w^j wzór (15).

4 Przestrzeń wektorowa macierzy

Oznaczmy przez $\mathbb{K}^m{}_n$ zbiór macierzy $m \times n$ o elementach z $\mathbb{K}$.

Zbiorowi $\mathbb{K}^m{}_n$ można nadać strukturę przestrzeni liniowej w następujący sposób: Dodawanie macierzy określamy jako:

$$\begin{bmatrix} a^1{}_1 & \dots & a^1{}_n \\ \vdots & & \vdots \\ a^m{}_1 & \dots & a^m{}_n \end{bmatrix} + \begin{bmatrix} b^1{}_1 & \dots & b^1{}_n \\ \vdots & & \vdots \\ b^m{}_1 & \dots & b^m{}_n \end{bmatrix} = \begin{bmatrix} a^1{}_1 + b^1{}_1 & \dots & a^1{}_n + b^1{}_n \\ \vdots & & \vdots \\ a^m{}_1 + b^m{}_1 & \dots & a^m{}_n + b^m{}_n \end{bmatrix},$$

a mnożenie macierzy przez liczbę jako:

$$\lambda \cdot \begin{bmatrix} a^1{}_1 & \dots & a^1{}_n \\ \vdots & & \vdots \\ a^m{}_1 & \dots & a^m{}_n \end{bmatrix} = \begin{bmatrix} \lambda a^1{}_1 & \dots & \lambda a^1{}_n \\ \vdots & & \vdots \\ \lambda a^m{}_1 & \dots & \lambda a^m{}_n \end{bmatrix}.$$

Takie zdefiniowanie dodawania macierzy i mnożenia ich przez liczbę zadaje izomorfizm pomiędzy $L(V, W)$ a $\mathbb{K}^m{}_n$.

5 Macierz złożenia odwzorowań

Niech U, V, W będą trzema przestrzeniami wektorowymi z wybranymi bazami: $g = (g_1, \dots, g_p)$ – baza w U; $e = (e_1, \dots, e_n)$ – baza w V; $f = (f_1, \dots, f_m)$ – baza w W. Jeśli $T \in L(V, W)$ i $S \in L(U, V)$, to $R = T \circ S \in L(U, W)$. Znajdźmy, jak macierz złożenia odwzorowań T i S:

$$[R]^f{}_g = C = \begin{bmatrix} c^1{}_1 & \dots & c^1{}_p \\ \vdots & & \vdots \\ c^m{}_1 & \dots & c^m{}_p \end{bmatrix}$$

wyraża się przez macierze tych odwzorowań:

$$[T]^f{}_e = A = \begin{bmatrix} a^1{}_1 & \dots & a^1{}_n \\ \vdots & & \vdots \\ a^m{}_1 & \dots & a^m{}_n \end{bmatrix}, \qquad [S]^e{}_g = B = \begin{bmatrix} b^1{}_1 & \dots & b^1{}_p \\ \vdots & & \vdots \\ b^n{}_1 & \dots & b^n{}_p \end{bmatrix}.$$

Współczynniki $c^i{}_j$ macierzy $C = [R]^f{}_g$ są określone przez: Jeśli $U \ni u = \sum_{j=1}^p u^j g_j$, $W \ni w = \sum_{i=1}^p u^i g_i$, to

$$w = Ru \qquad \Longleftrightarrow \qquad w^i = \sum_{j=1}^p c^i{}_j u^j.$$

Ponieważ $w = Tv$, gdzie $v = Su$, to oznaczając przez v^k współrzędne wektora v w bazie e (mamy więc $v = \sum_{k=1}^n v^k e_k$) mamy

$$w^i = \sum_{k=1}^n a^i{}_k v^k = \sum_{k=1}^n a^i{}_k \left(\sum_{j=1}^p b^k{}_j u^j \right) = \sum_{j=1}^p \left(\sum_{k=1}^n a^i{}_k b^k{}_j \right) u^j.$$

Współczynniki $c^i{}_j$ są więc dane przez

$$c^i{}_j = \sum_{k=1}^n a^i{}_k b^k{}_j. \tag{16}$$

Def. Jeśli $A \in \mathbb{K}^m{}_n$, $B \in \mathbb{K}^n{}_p$, to macierz $C \in \mathbb{K}^m{}_p$, której elementy $c^i{}_j$ dane są przez wzór (16), nazywamy *iloczynem macierzy* A i B (i piszemy: $C = AB$).

Podsumowując, *macierz złożenia odwzorowań jest iloczynem macierzy tych odwzorowań*:

$$[T \circ S]^f{}_g = [T]^f{}_e [S]^e{}_g.$$

5.1 Własności mnożenia macierzy

Mnożenie macierzy jest *łączne*, tzn. $(AB)C = A(BC)$, jeśli iloczyny AB i BC są określone z sensem (łączność mnożenia macierzy wynika z łączności składania odwzorowań). Natomiast mnożenie macierzy na ogół nie jest przemienne:

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Jeśli przez Id_V oznaczymy odwzorowanie identycznościowe $V \to V$, to dla $T \in L(V, W)$ mamy

$$[T]^f{}_e = [T \circ \mathrm{Id}_V]^f{}_e = [T]^f{}_e [\mathrm{Id}_V]^e{}_e.$$

Podobnie

$$[T]^f{}_e = [\mathrm{Id}_W \circ T]^f{}_e = [\mathrm{Id}_W]^f{}_f [T]^f{}_e$$

dla Id_W – odwzorowania identycznościowego $W \to W$.

Macierz

$$[\mathrm{Id}_V]^e{}_e = ([(\mathrm{Id}_V)e_1]^e, \dots, [(\mathrm{Id}_V)e_n]^e) = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \overset{ozn}{=} I_n$$

nazywamy *macierzą jednostkową*. Pomnożenie dowolnej macierzy A przez macierz jednostkową (w przypadkach, kiedy jest to wykonalne) nie zmienia A (własność elementu neutralnego): Dla $A \in \mathbb{K}^m{}_n$ mamy

$$A\, I_n = A = I_m\, A.$$

Def. Mówimy, że macierz $A \in \mathbb{K}^n{}_n$ jest *odwracalna*, jeśli istnieje macierz $B \in \mathbb{K}^n{}_n$ taka, że $AB = I_n = BA$. W takim przypadku B nazywamy macierzą *odwrotną* do A.

Łatwo zobaczyć, że powyższa macierz B, jeśli istnieje, jest określona jednoznacznie. Zapisujemy ją jako A^{-1}. Zachodzi: $(A^{-1})^{-1} = A$.

Gdy mamy odwzorowania (odwracalne) S, T, to odwzorowanie odwrotne do $S \circ T$ jest: $(S \circ T)^{-1} = T^{-1} \circ S^{-1}$. To daje wzór na macierz odwrotną do iloczynu: Jeśli A, C – macierze (odwracalne), to $(AC)^{-1} = C^{-1}A^{-1}$.

Macierze izomorfizmów (tzn. odwzorowań liniowych odwracalnych) są odwracalne: Jeśli $T \in L(V, W)$ jest odwracalne, e – baza w V, f – baza w W, to

$$[T]^f{}_e [T^{-1}]^e{}_f = [T \circ T^{-1}]^f{}_f = [\mathrm{Id}_W]^f{}_f = I_n, \qquad [T^{-1}]^e{}_f [T]^f{}_e = [T^{-1} \circ T]^e{}_e = [\mathrm{Id}_V]^e{}_e = I_n,$$

gdzie $n = \dim V = \dim W$.

5.2 Zmiana baz i zmiana macierzy operatora przy zmianie bazy

Jeśli $T \in L(V,W)$ oraz e, e' – dwie bazy w V, zaś f, f' – dwie bazy w W, to

$$[T]^{f'}{}_{e'} = [\mathrm{Id}_W \circ T \circ \mathrm{Id}_V]^{f'}{}_{e'} = [\mathrm{Id}_W]^{f'}{}_f [T]^f{}_e [\mathrm{Id}_V]^e{}_{e'}.$$

Wzór ten pozwala obliczyć macierz odwzorowania T w "‘nowych"’ bazach e', f' mając macierz T w "‘starych"’ bazach e, f.

Macierz

$$[\mathrm{Id}_V]^e{}_{e'} = ([e'_1]^e, \ldots, [e'_n]^e)$$

nazywa się *macierzą zmiany bazy*. Macierz $[\mathrm{Id}_V]^e{}_{e'}$ jest macierzą odwrotną do $[\mathrm{Id}_V]^{e'}{}_e$:

$$[\mathrm{Id}_V]^e{}_{e'} [\mathrm{Id}_V]^{e'}{}_e = [\mathrm{Id}_V \circ \mathrm{Id}_V]^e{}_e = [\mathrm{Id}_V]^e{}_e = I_n,$$

i analogicznie

$$[\mathrm{Id}_V]^{e'}{}_e [\mathrm{Id}_V]^e{}_{e'} = [\mathrm{Id}_V \circ \mathrm{Id}_V]^{e'}{}_{e'} = [\mathrm{Id}_V]^{e'}{}_{e'} = I_n.$$

Przykład. Niech operator A w bazie standardowej $e = (e_1, e_2) = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)$ będzie dany macierzą:

$$[A]^e{}_e = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}.$$

Znajdźmy jego macierz w bazie $e' = (e'_1, e'_2) = \left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right)$

Mamy:

$$[\mathrm{Id}]^e{}_{e'} = ([e'_1]^e, [e'_2]^e) = \begin{bmatrix} 2 & 2 \\ 1 & -1 \end{bmatrix}$$

(poszczególne kolumny są to składowe wektorów e'_i w bazie wektorów e_k). Macierz $[\mathrm{Id}]^{e'}{}_e$ jest macierzą *odwrotną* do $[\mathrm{Id}]^e{}_{e'}$. Niedługo zobaczymy, jak się liczy macierz odwrotną. W tym momencie niech Czytelnik sprawdzi, że $[\mathrm{Id}]^{e'}{}_e = \frac{1}{4} \begin{bmatrix} 1 & 2 \\ 1 & -2 \end{bmatrix}$ tzn. $[\mathrm{Id}]^{e'}{}_e [\mathrm{Id}]^e{}_{e'} = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Mamy

$$[A]^{e'}{}_{e'} = [\mathrm{Id}]^{e'}{}_e [A]^e{}_e [\mathrm{Id}]^e{}_{e'} = \frac{1}{4} \begin{bmatrix} 1 & 2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$$

W tym przypadku, przez odpowiedni wybór bazy, udało się wyrazić operator A w postaci *diagonalnej*; taka postać jest bardzo ważna przy badaniu macierzy.

5.3 Układy równań liniowych

Układ równości

$$\left\{ \begin{array}{ccccccc} a^1{}_1 x^1 & + & \ldots & + & a^1{}_n x^n & = & b^1 \\ a^2{}_1 x^1 & + & \ldots & + & a^2{}_n x^n & = & b^2 \\ \vdots & & & & \vdots & & \vdots \\ a^m{}_1 x^1 & + & \ldots & + & a^m{}_n x^n & = & b^m \end{array} \right. \tag{17}$$

gdzie $a^i{}_j$, b^i – dane liczby (z ciała $\mathbb{K}$), zaś $x^1, \ldots, x^n$ – liczby niewiadome, nazywa się *układem m równań liniowych z n niewiadomymi.*

Przepiszmy lewą stronę układu równań na trzy sposoby:

$$x^1 \begin{bmatrix} a^1{}_1 \\ \vdots \\ a^m{}_1 \end{bmatrix} + \dots + x^n \begin{bmatrix} a^1{}_n \\ \vdots \\ a^m{}_n \end{bmatrix} = \begin{bmatrix} a^1{}_1x^1 & + & \dots & + & a^1{}_nx^n \\ a^2{}_1x^1 & + & \dots & + & a^2{}_nx^n \\ \vdots & & & & \vdots \\ a^m{}_1x^1 & + & \dots & + & a^m{}_nx^n \end{bmatrix} \tag{18}$$

$$= \begin{bmatrix} a^1{}_1 & \dots & a^1{}_n \\ a^2{}_1 & \dots & a^2{}_n \\ \vdots & & \vdots \\ a^m{}_1 & \dots & a^m{}_n \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix}.$$

Stąd, wprowadzając oznaczenia:

$$a_1 = \begin{bmatrix} a^1{}_1 \\ a^2{}_1 \\ \vdots \\ a^m{}_1 \end{bmatrix}, \ \dots, \ a_n = \begin{bmatrix} a^1{}_n \\ a^2{}_n \\ \vdots \\ a^m{}_n \end{bmatrix}, \quad A = (a_1, \dots, a_n) = \begin{bmatrix} a^1{}_1 & \dots & a^1{}_n \\ a^2{}_1 & \dots & a^2{}_n \\ \vdots & & \vdots \\ a^m{}_1 & \dots & a^m{}_n \end{bmatrix}$$

oraz

$$x = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix}, \quad b = \begin{bmatrix} b^1 \\ b^2 \\ \vdots \\ b^m \end{bmatrix}$$

układ równań (17) można zapisać w postaci

$$x^1a_1 + \dots + x^na_n = b$$

lub

$$Ax = b.$$

Macierz A nazywa się *macierzą układu równań* (17).

Można rozpatrywać ją jako macierz pewnego operatora liniowego $\mathbb{K}^n \to \mathbb{K}^m$.

Ze względu na lewą część (18) mamy

$$\langle a_1, \dots, a_n \rangle = \mathrm{Im} A.$$

Def. Liczbę

$$\mathrm{rank}(A) = \dim \mathrm{Im}\, A$$

nazywamy *rzędem* (kolumnowym) macierzy A. Jest oczywiste, że jest to też ilość liniowo niezależnych kolumn macierzy A.

Zachodzi następujące stwierdzenie dotyczące istnienia rozwiązań układu (17).

Stwierdzenie. Następujące warunki są równoważne:

1. Istnieje rozwiązanie układu (17).
2. $b \in \mathrm{Im}\, A$.
3. $b \in \langle a_1, \dots, a_n \rangle$.

4. $\langle a_1, \ldots, a_n, b\rangle = \langle a_1, \ldots, a_n\rangle$

5. $\operatorname{rank}(A, b) = \operatorname{rank}(A)$

Tutaj $(A, b) = (a_1, \ldots, a_n, b)$ oznacza macierz powstałą z A przez dołączenie jeszcze jednej kolumny – wektora b. Macierz (A, b) nazywa się *macierzą rozszerzoną* układu (17). Zawiera ona pełną informację o układzie.

Uwagi.

- Równoważność 1) i 5) nosi nazwę *twierdzenia Kroneckera – Capelliego.*
- Jeśli prawa strona układu jest zerem (tzn. $b = \mathbf{0}$), to układ nazywamy *jednorodnym*; jeśli $b \neq \mathbf{0}$), to układ nazywamy *niejednorodnym.*
- Zbiór rozwiązań układu jednorodnego jest przestrzenią wektorową (podprzestrzenią $\mathbb{K}^n$). Zbiór rozwiązań układu niejednorodnego *nie* jest przestrzenią wektorową.

Mamy też proste

Stwierdzenie. Jeśli $m = n$, to równanie $Ax = b$ ma dla każdego b dokładnie jedno rozwiązanie wtedy i tylko wtedy, gdy istnieje A^{-1}. Wówczas $x = A^{-1}b$.

W praktyce rzadko rozwiązuje się układy równań przez znajdowanie macierzy odwrotnej. ("'Rzadko"' nie znaczy "'wcale"'; jest to sposób potrzebny gdy np. macierz A zależy od parametrów). Najczęściej robi się to metodą *redukcji wierszowej* – dodawania równań stronami tak, by otrzymać możliwie prostą macierz układu. Wykorzystujemy tu oczywisty fakt, że dodając do jakiegoś równania kombinację liniową pozostałych równań, otrzymujemy układ równoważny, tzn. mający te same rozwiązania. Operacja ta odpowiada przejściu do innej bazy w przestrzeni W.

Z formalnego punktu widzenia, redukcja wierszowa jest kombinacją następujących *operacji elementarnych*, z których każda w oczywisty sposób nie zmienia rozwiązań układu:

1. Mnożenie wiersza przez niezerową stałą.
2. Dodawanie jednego wiersza do drugiego.
3. Przestawianie dwóch wierszy.

Przykład.
1) Rozwiązać układ równań:

$$\begin{bmatrix} 3 & -4 & 1 \\ -5 & 2 & 3 \\ 2 & 1 & -3 \end{bmatrix} x = \begin{bmatrix} 2 \\ -8 \\ 5 \end{bmatrix}$$

Piszemy rozszerzoną macierz układu i redukujemy (wierszowo).

$$\left[\begin{array}{ccc|c} 3 & -4 & 1 & 2 \\ -5 & 2 & 3 & -8 \\ 2 & \boxed{1} & -3 & 5 \end{array}\right] \overset{(1)}{\sim} \left[\begin{array}{ccc|c} 11 & 0 & -11 & 22 \\ -9 & 0 & 9 & -18 \\ 2 & 1 & -3 & 5 \end{array}\right] \overset{(2)}{\sim} \left[\begin{array}{ccc|c} -1 & 0 & 1 & -2 \\ -1 & 0 & 1 & -2 \\ 2 & 1 & -3 & 5 \end{array}\right]$$

Równoważność (1) bierze się stąd, że równanie trzecie dodawaliśmy do pierwszego i drugiego z takimi współczynnikami, żeby w drugiej kolumnie pojawiły się same zera. Równoważność (2): Pomnożyliśmy równania: pierwsze i drugie przez odpowiednio 11 i -9, aby otrzymać prostszą postać. I dalej:

$$\overset{(3)}{\sim} \left[\begin{array}{ccc|c} -1 & 0 & \boxed{1} & -2 \\ 2 & 1 & -3 & 5 \end{array}\right] \overset{(4)}{\sim} \left[\begin{array}{ccc|c} -1 & 0 & 1 & -2 \\ -1 & 1 & 0 & -1 \end{array}\right] \overset{(5)}{\sim} \left[\begin{array}{ccc|c} -1 & 1 & 0 & -1 \\ -1 & 0 & 1 & -2 \end{array}\right]$$

Równoważność (3): Skreśliliśmy jedno z powtarzających się równań. (4): Dodaliśmy pierwsze równanie do drugiego z takim współczynnikiem, by wyzerować w nim trzecią kolumnę. (5): Przestawiliśmy wiersze, aby uzyskać jawnie diagonalną podmacierz układu.

Ostatnia macierz układu oznacza, że wyjściowy układ doprowadziliśmy do postaci

$$\begin{cases} -x_1 & + & x_2 & = & -1 \\ -x_1 & + & x_3 & = & -2 \end{cases}$$

którego rozwiązanie od razu piszemy:

$$\begin{cases} x_2 & = & x_1 - 1 \\ x_3 & = & x_1 - 2 \end{cases}$$

lub, w równoważnej postaci (zmieniając jeszcze nazwę x_1 na λ)

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ -2 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

2) Układ równań:

$$\begin{bmatrix} 3 & -4 & 1 \\ -5 & 2 & 3 \\ 2 & 1 & -3 \end{bmatrix} x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

jest sprzeczny.

3) Będzie tu jeszcze liczenie macierzy odwrotnej.

Metoda redukcji wierszowej (i kolumnowej, gdzie analogiczne operacje przeprowadza się na kolumnach macierzy) mają szerokie zastosowanie jako metody rachunkowe do znajdowania jądra i obrazu operatora czy macierzy odwrotnej.

W analogii do rzędu kolumnowego, definiuje się też *rząd wierszowy*:

Def. Rząd wierszowy macierzy A: $\mathrm{rank}_w(A)$ to ilość liniowo niezależnych wierszy tej macierzy.

Okazuje się, że zachodzi *równość* tych dwu rzędów:

Tw. $\mathrm{rank}_w(A) = \mathrm{rank}_k(A)$.

Dowód. Rozpatrzmy najsampierw jednorodny układ równań liniowych:

$$\begin{cases} a^1{}_1 x^1 & + & a^1{}_2 x^2 & + & \dots & + & a^1{}_n x^n & = & 0 \\ a^2{}_1 x^1 & + & a^2{}_2 x^2 & + & \dots & + & a^2{}_n x^n & = & 0 \\ \vdots & & & & & & \vdots & & \vdots \\ a^m{}_1 x^1 & + & a^m{}_2 x^2 & + & \dots & + & a^m{}_n x^n & = & 0 \end{cases} \tag{19}$$

Dow. pomijamy.